

修 士 論 文 の 和 文 要 旨

大学院	電気通信学研究科	博士前期課程	情報通信工学専攻
氏 名	佐藤 めぐみ	学籍番号	0430017
論 文 題 目	HMM-GMM合成とマルチパスモデルを用いた雑音環境下での音声認識		
<p>要 旨</p> <p>音声は多くの人がごく自然に使うことのできるメディアの一つである。そこで現在、様々な場面において音声認識技術の実用化が始まっており、更にその応用分野を拡大していくことが期待されている。しかしまだ音声認識技術は完全なものではなく問題は残されている。本研究ではそのうち2つの課題に着目した。</p> <p>1つ目の課題は、環境雑音に対して頑健なシステムを構築することである。雑音は本来の音声の特性を変化させる為、認識対象の音声とシステムが持つ音声モデルとの間に差異を生じさせ、誤認識をしてしまう原因となりうる。よって、雑音環境下においても対象音声の音響特性に精密に近似したモデルを作成することが重要である。この差異を生じさせる雑音混入音声の性質の一つに、時間領域上におけるSNR(Signal to Noise Ratio)の非定常性が挙げられる。この対処法としてマルチSNR-マルチパス法の有効性が確認されている。これは、同じ音声に複数のSNRで雑音の要素を付加し、それぞれ独立にモデルを作成し並列に並べた後、適切なパスを選択しながら全体の尤度を計算する手法である。</p> <p>2つ目の課題は、音声認識の対象を不特定話者として認識システムの汎用性を高めることである。そこで本研究では音声認識システム「Julius」が持つ「高精度成人モデル」をベースに用いることにした。Juliusが持つ音声モデルは、クリーンな環境においては不特定話者に対して高い精度を見せる。このJuliusモデルに対して雑音の要素を付加する方法としてHMM-GMM合成法を用いた。これは、音声に雑音を混入した後にモデル化する手法とは異なり、クリーンな音声と雑音を独立に学習した後モデルの段階で合成する手法である。</p> <p>本研究ではHMM-GMM合成法とマルチSNR-マルチパス法との組合せによって、より汎用性が高く精密なモデルの作成を目指した。</p> <p>まず、雑音GMMはどのような構造にしたら認識性能の向上に貢献できるのかを検討した。次に合成に必要なパラメータの一部がJulius-HMMには欠けている為、実験的にこれを設定した。これらを踏まえてパラメータを設定し合成モデルを作成して、複数話者に対して雑音環境下における単語音声認識実験を行ったところ、クリーンモデルを用いた場合に比べて最高で約15ポイント認識率が向上した。今後、より良く雑音の特性を反映する合成法を検討することにより、さらに認識率が向上する可能性がある。</p>			